8 February 2011

Sandia National Laboratories

HPC and UQ Capabilities in Center 1400 Relevant to Smart Grid Technologies

Jim Kamm & Tim Trucano

Optimization and Uncertainty Quantification Dept., Org 1441

SAND2011-0954P

Sandia National Laboratories is a multi program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

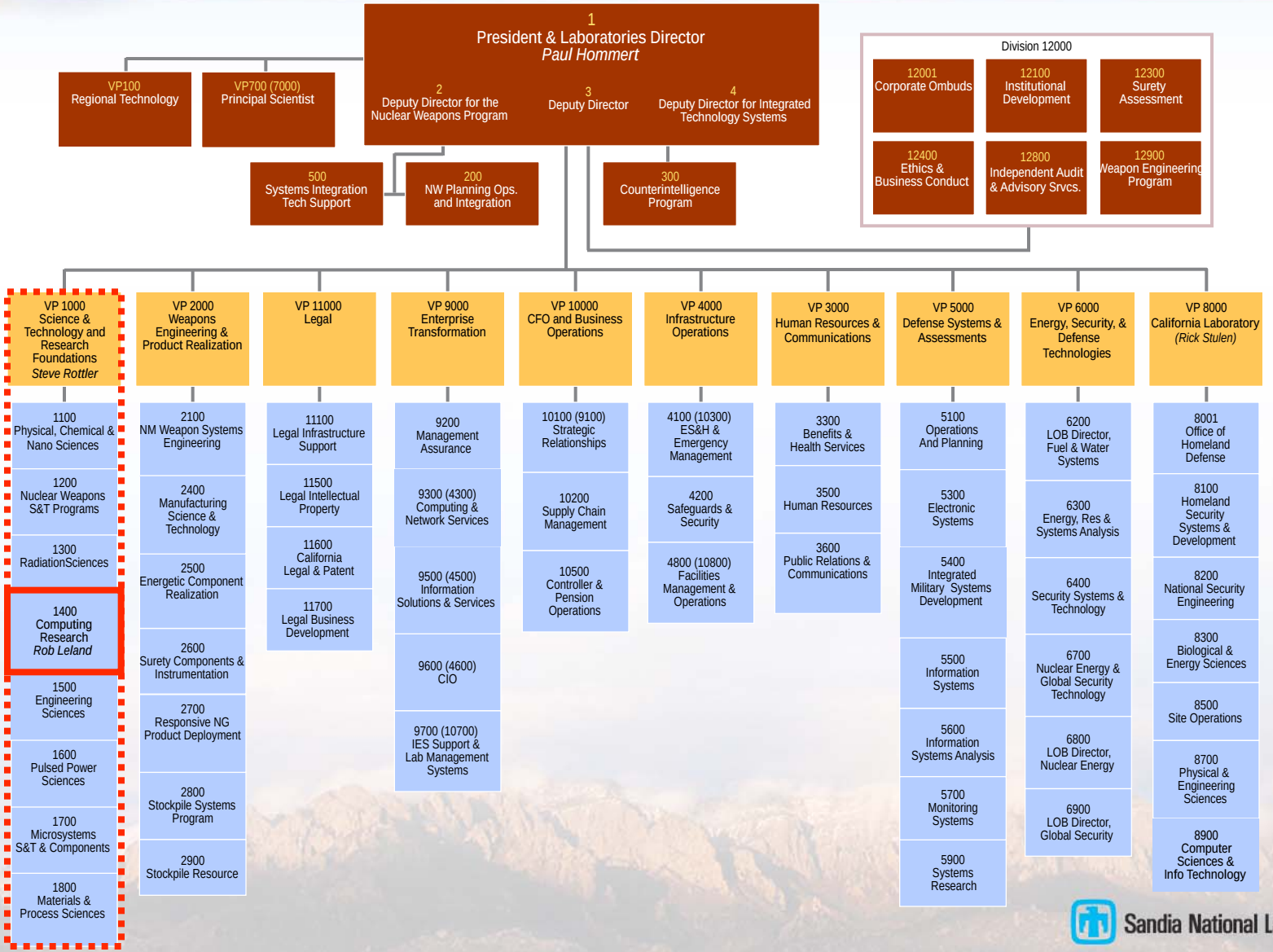
.

Sandia National Laboratories

Brief background to help calibrate you…

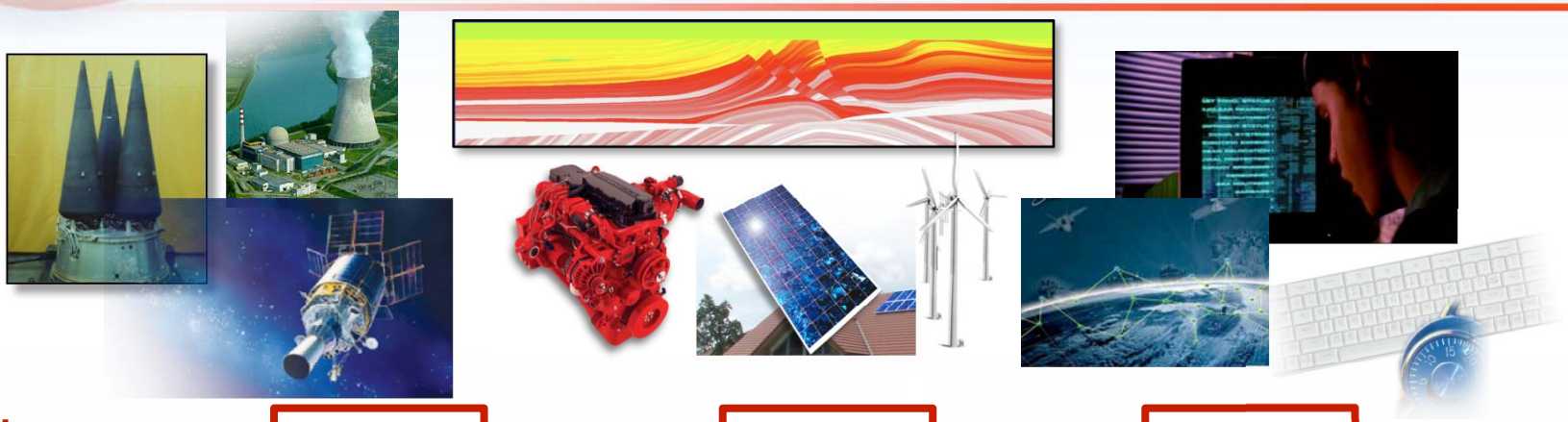
- **Sandia Computing Research Center (Org. "1400")**
  - **Director: Rob Leland**, PhD Parallel Computing
  - **Mission:** Deliver computer science, mathematical science, computational science, and information science capabilities, tools, and platforms that enable Sandia's mission.
  - **Goal:** Drive transformation through predictive simulation.
- **Jim Kamm** jrkamm@sandia.gov
  - PhD Applied Mathematics
  - ~25 year SAIC/LANL/Sandia career
  - Computational compressible flow, applied math, and V&V/UQ
- **Tim Trucano** tgtruca@sandia.gov
  - PhD Mathematics
  - 30+ years Sandia
  - Mathematical physics, computational shock wave physics, and V&V/UQ

Where is Center 1400 in Sandia's organization structure?



Sandia National Laboratories

Sandia occupies a national position in computational science supporting predictive modeling and simulation.



Sandia Strategic Thrusts

- Nuclear Security
- Energy Security
- Cyber Security

Computer Science Application Themes

- Nuclear Weapons Engineering
- Energy/Climate Security
- Cyber Assurance
- *Decision Support*

Computer Science Technology Themes

- Predictive Simulation
- Scalable Computing
- Scalable Informatics
- Cognitive Science
- *Enabling Technologies*

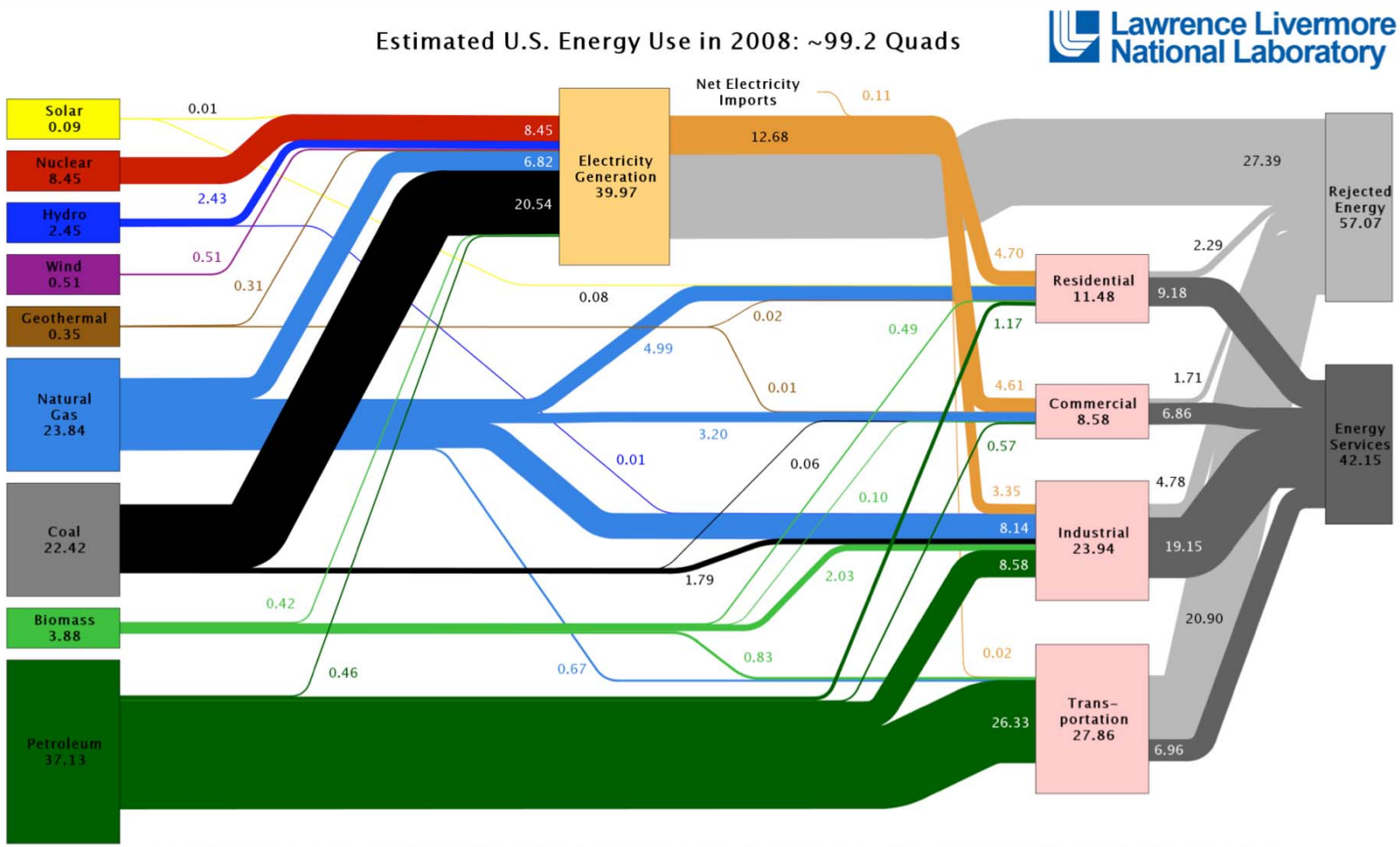
Sandia National Laboratories

DOE mission imperatives require simulation and analysis for policy and decision making.

- ***Climate Change*: Understanding, mitigating and adapting to the effects of global warming**
  - Sea level rise
  - Severe weather
  - Regional climate change
  - Geologic carbon sequestration
- ***Energy*: Reducing U.S. reliance on foreign energy sources and reducing the carbon footprint of energy production**
  - Reducing time and cost of reactor design and deployment
  - Improving the efficiency of combustion energy systems
- ***National Nuclear Security*: Maintaining a safe, secure and reliable nuclear stockpile**
  - Stockpile certification
  - Predictive scientific challenges
  - Real-time evaluation of urban nuclear detonation

Accomplishing these missions requires exascale resources.

US energy flows (2008): ≈ 104 Exajoules



Sandia National Laboratories

Sandia Center 1400 has specific capabilities relevant to the Sandia/UVM collaboration.

Center 1400 has expertise in:

- High Performance Computing (HPC)
- Uncertainty Quantification (UQ), V&V, Optimization
- Cognitive Science

What I represent from Sandia

Fundamental to our research portfolio, we strive to:

- Develop innovations in methods and algorithms
- Be motivated by applications
- Help drive development of HPC platforms
- Employ ubiquitous V&V and UQ

The vision is high-impact utilization of computing sciences in problems having societal consequences.

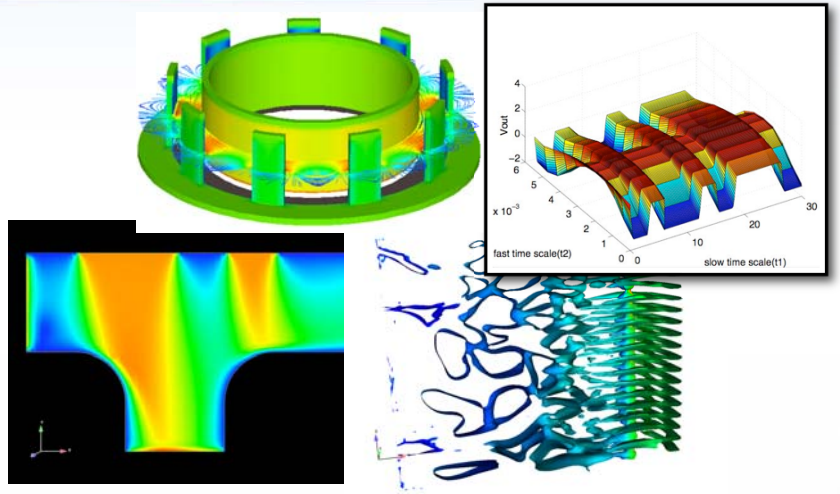
Multidisciplinary research is key to the mission of Center 1400.

Trilinos

http://trilinos.sandia.gov/

DAKOTA

http://dakota.sandia.gov/



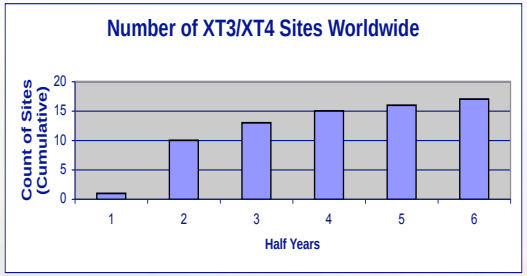
Leading Edge Algorithms and Enabling Technologies

State-of-Art Computational Science and Visualization Applications

CSRI

Computer Science Research Institute

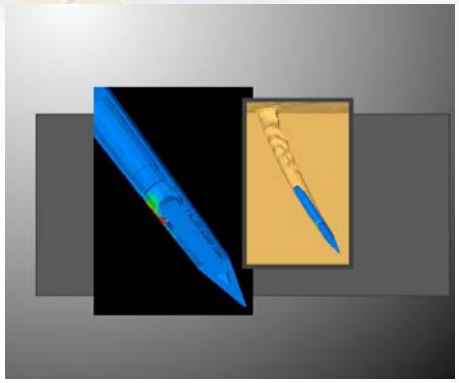
http://www.cs.sandia.gov/CSRI/

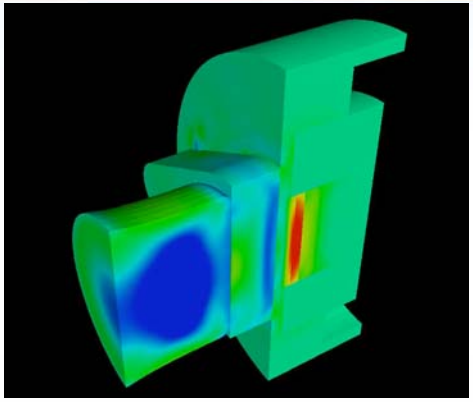


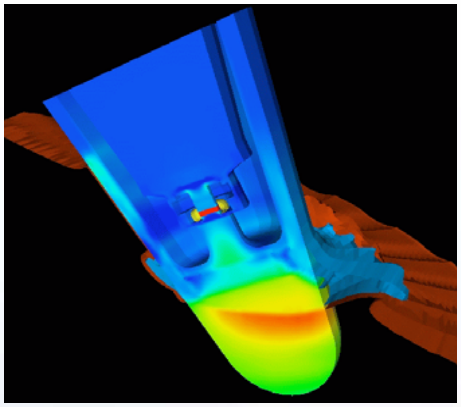
Scalable HPC Architectures and Systems

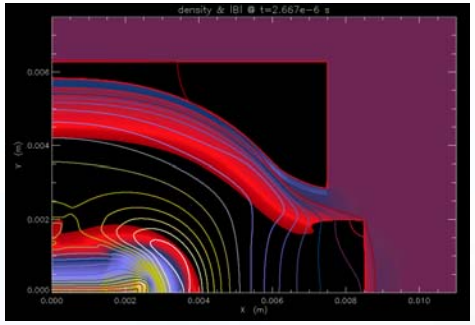
Sandia National Laboratories

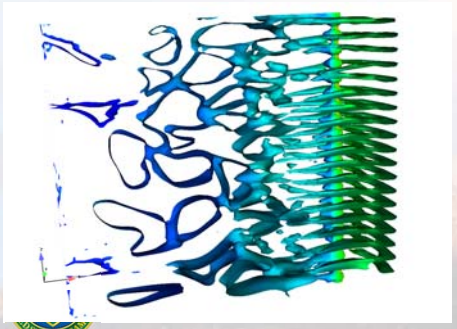
ALEGRA is a multi-physics simulation code with diverse M&S capabilities.

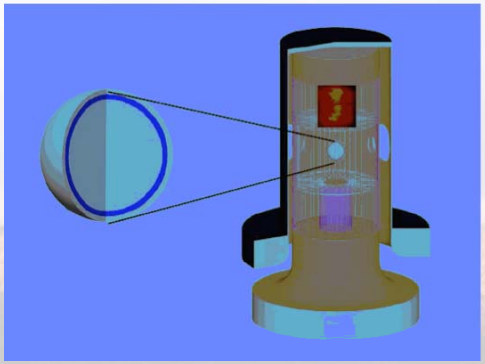












HEDP Theory and ICF Target Design

Overview

- The ALEGRA suite of applications model shock and high energy environments for solids, fluids and plasmas using a multi-material arbitrary Lagrangian-Eulerian (ALE) multi-physics methodology.
- The ALEGRA applications run on large parallel message passing architectures in 2D and 3D geometries.

ALEGRA

- Armor design and Analysis
- Shaped Charge & EFP performance
- Railgun & Coilgun Design and Analysis

ALEGRA-HEDP

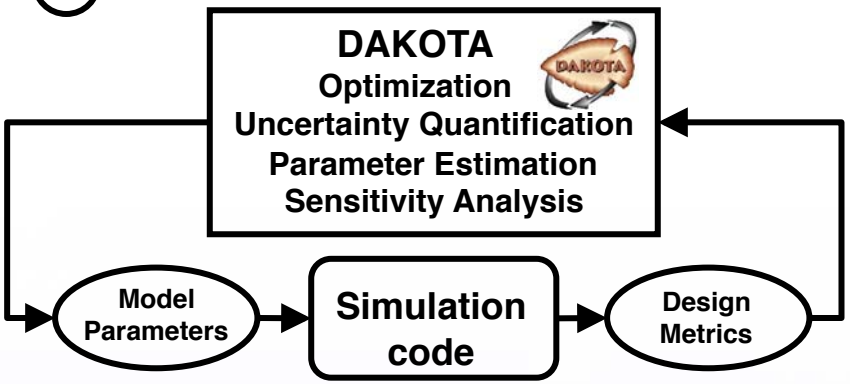
- Magnetohydrodynamics
- Z-pinch
- Power Flow
- ICE/Magnetic Flyers
- ICF

Computational Shock and Multiphysics

Sandia National Laboratories

There are specific Sandia activities that share common challenges with this collaboration.

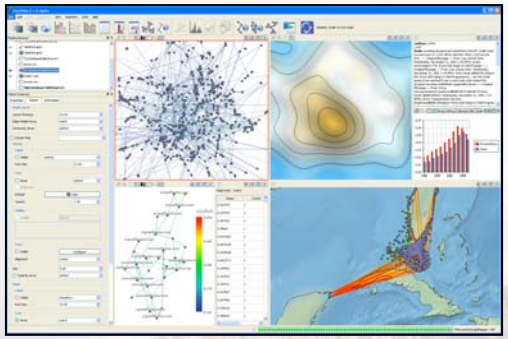
UQ tools developed at Sandia have enabled many advances. ①



Sandia is a leader in development, implementation, and utilization of HPC technology. ②

Red Storm cielo *Exascale*

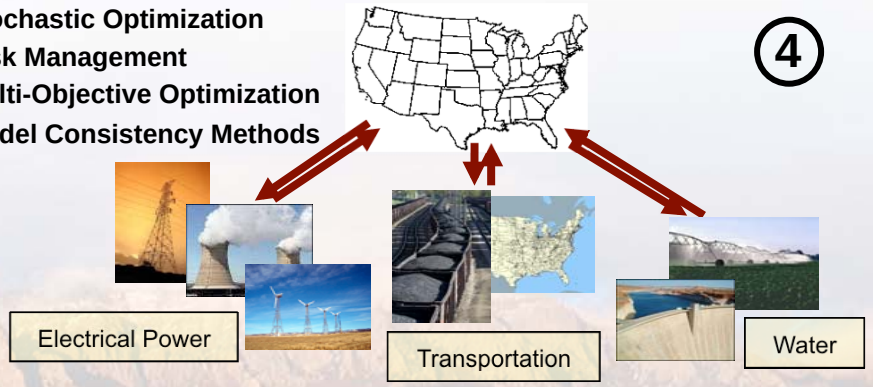
Scalable SciViz and InfoViz tools for analysis of complex simulation data ③



Technology is being developed for informatics visualization to promote understanding of complex data.

Optimization of Large-Scale Heterogeneous System-of-Systems ④

- Stochastic Optimization
- Risk Management
- Multi-Objective Optimization
- Model Consistency Methods



Sandia National Laboratories

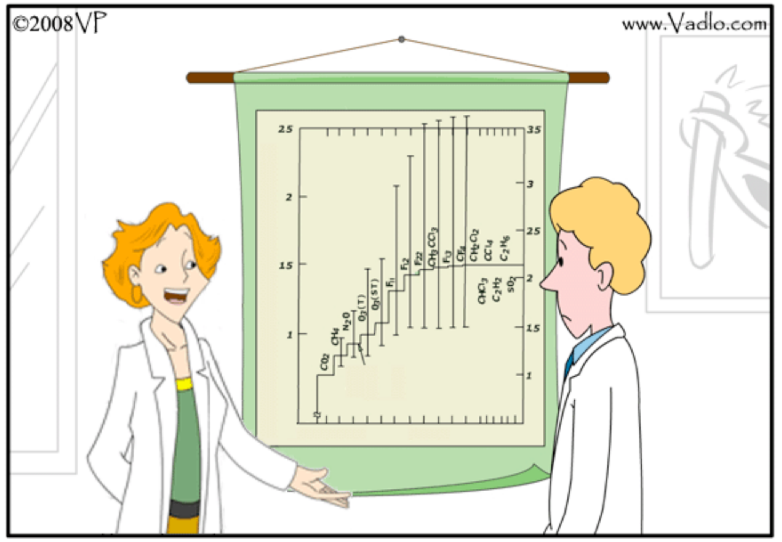
① "Admiral's Slide" summary of UQ

- **UQ in computational science is the formal characterization, propagation, aggregation, comprehension, and communication of aleatory (variability) and epistemic (incomplete knowledge) uncertainties.**
  - E.g., demand fluctuations in a power grid are *aleatory uncertainties*.
  - E.g., incomplete knowledge about the future (scenarios uncertainty), the validity of models, and inadequate "statistics" are *epistemic uncertainties*.
  - E.g., sensitivity analysis is ONE procedure under the overall UQ umbrella — it helps drive parsimony.
- **A huge range of technical issues arise in the M&S components of problem definition and execution phases.**
- **Another huge range of technical issues (partially disjoint) arises in the delivery phase, especially in high-risk decision environments.**
- **"Probability" (including both frequency-based and subjective interpretations) is the main foundation for current "quantification."**
- **More complex epistemic uncertainties, for example arising in human interaction modeling, lead to other quantification formalisms (evidence theory, fuzzy sets, info-gap methods, etc.)**
- ***Any large-scale computational problem's computing requirements increase (usually significantly) with UQ.***

What is Uncertainty Quantification (UQ)?

- **Broadly speaking, UQ seeks to gauge the effect of system/model uncertainties on the observed/ computed outputs.**

- **The execution of UQ in M&S and the delivery of "prediction" typically has two distinct components:**
  - Characterization of uncertainty, typically quantitative characterizations for physical science M&S.
  - Reduction of uncertainty for purposes of improving prediction "accuracy."
  - For example, "error bars" in numerical simulations are ONE element of characterized uncertainty.



Do you really have to show the error bars?

Yes!

Uncertainty comes in two categories: Epistemic and Aleatory.

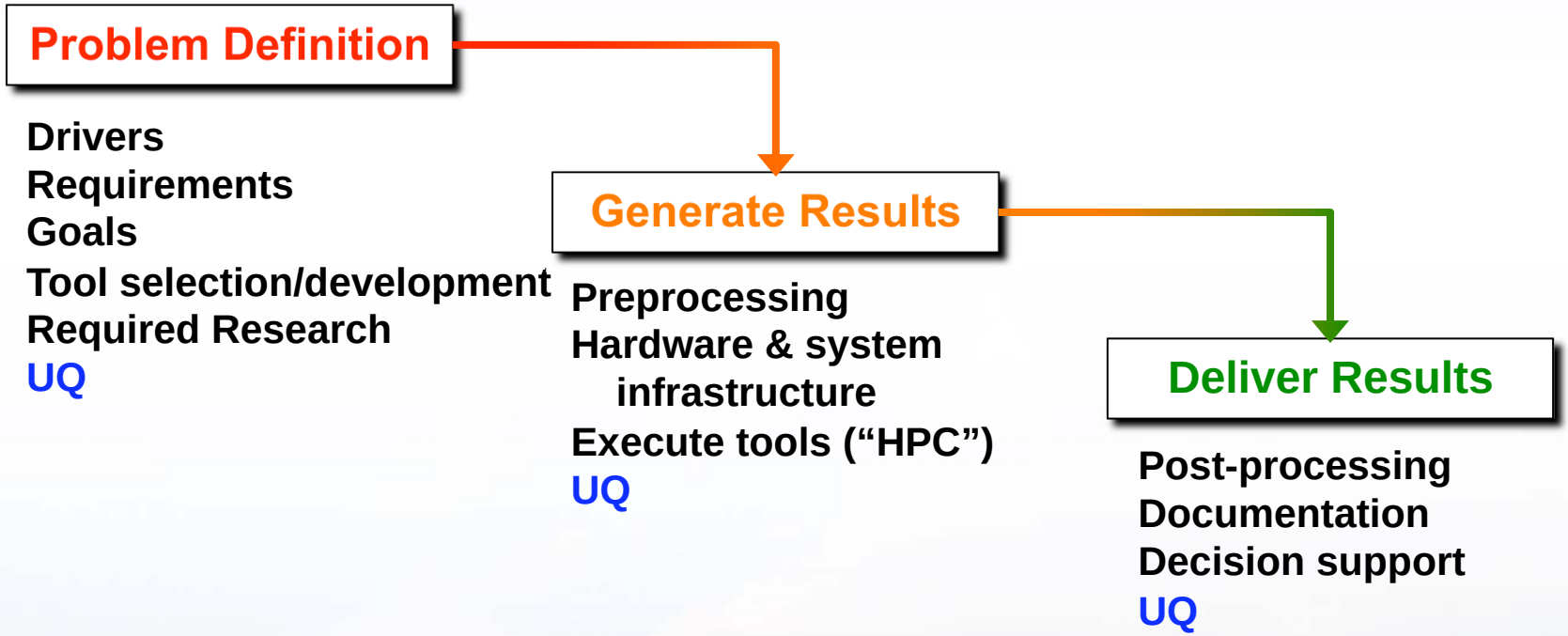
Epistemic Uncertainty

- Epistemology = "what distinguishes justified belief from opinion."
- Lack of knowledge about, say, the appropriate value to use for a quantity, or the proper model form to use.
- "Reducible uncertainty": can be reduced through increased understanding (research) or more, relevant data.
- E.g., the elastic modulus for the material in a specific component is presumably fixed but unknown or poorly known.
- Epistemic quantities have a fixed value in an analysis, but we do not know that fixed value.

Aleatory Uncertainty

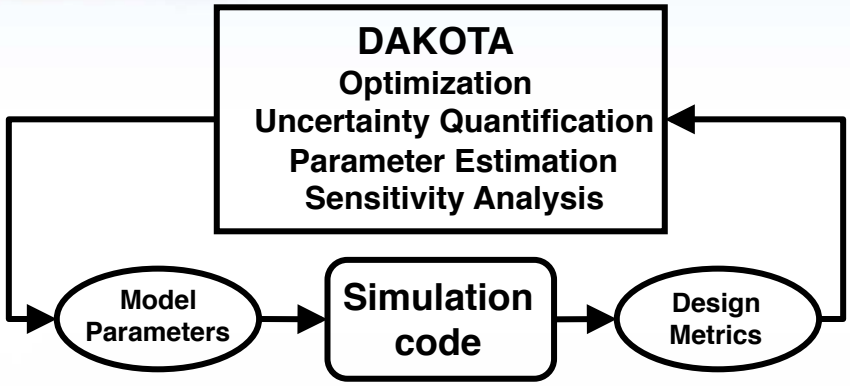
- "*Alea*" = Latin for "die"; Latin *aleator* = "dice player."
- Inherent randomness, intrinsic variability.
- "Irreducible uncertainty": cannot be reduced by additional data.
- E.g., the height of individuals in a population.
- Usually modeled with probability distributions.

Uncertainty Quantification (UQ) in the M&S workflow – a simplified view.

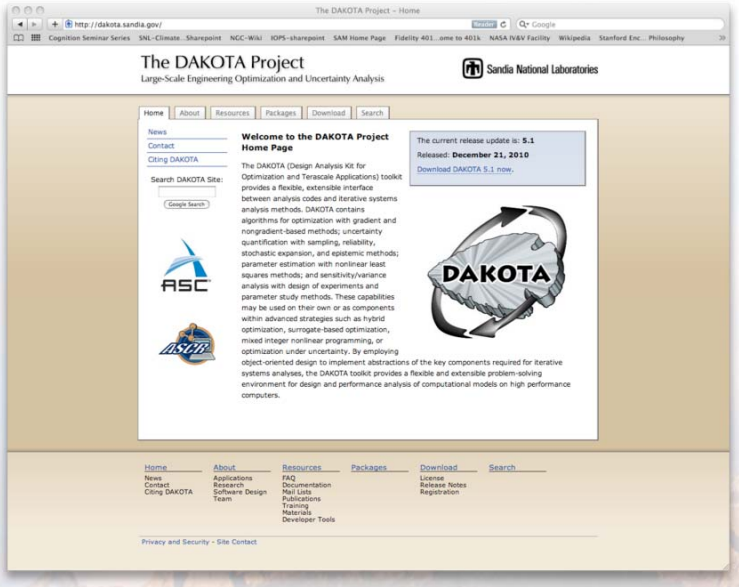


- **The presence of acknowledged uncertainty is fundamental — and fundamentally challenging.**
  - It complicates all aspects of the computational science that are required to address the problem.
  - V&V is just as ubiquitous!

Sandia's DAKOTA optimization and UQ toolkit is one software package.



- DAKOTA is openly available:

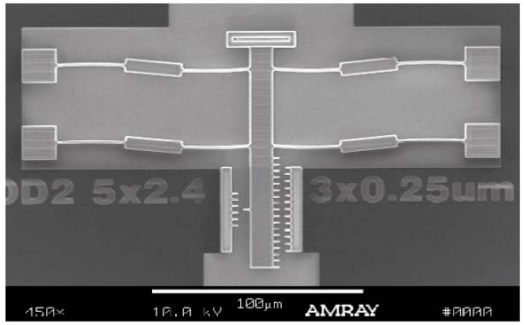


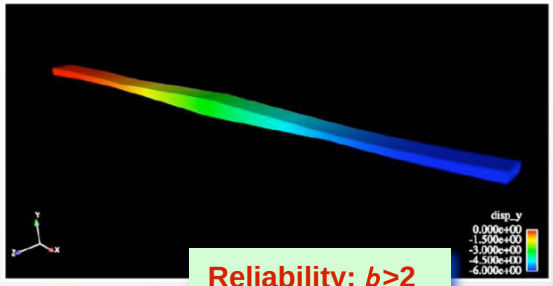
http://dakota.sandia.gov/

- Design and Analysis toolKit for Optimization and Terascale Applications includes a wide array of algorithm capabilities to support engineering transformation through advanced modeling and simulation.
- Optimization and UQ via DAKOTA has enabled many advances.

Optimization and UQ continue to increase in importance to M&S and its impact.

- **Problem: MEMS are subject to substantial variabilities**
  - Material properties, manufactured geometry, residual stresses
  - Part yields can be low or have poor durability
  - Data can be obtained → aleatory UQ → probabilistic methods
- **Goal: account for both uncertainties and errors in design**
  - Integrate UQ/OUU (DAKOTA), ZZ/QOI error estimation (Encore), adaptivity (SIERRA), nonlinear mechanics (Aria).
  - Perform solution verification in automated, parameter-adaptive way.
  - Generate fully converged UQ/OUU results at lower cost.
- **Results: optimized beam shape for bistable switch**
  - Achieved prescribed reliability & minimized sensitivity to uncertainties (robustness).

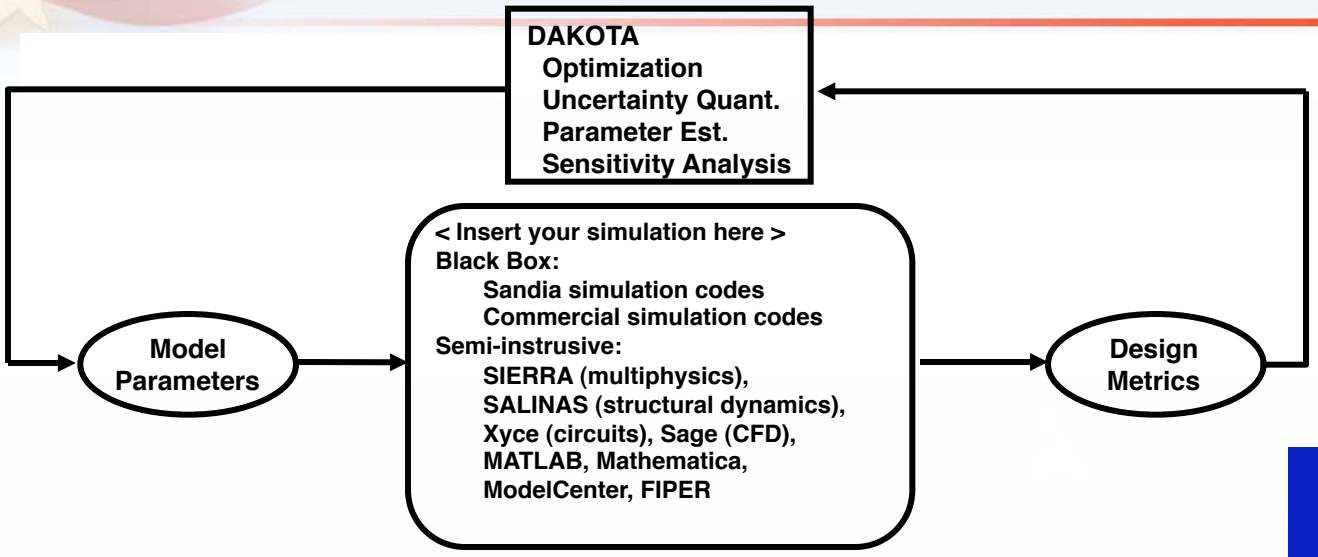


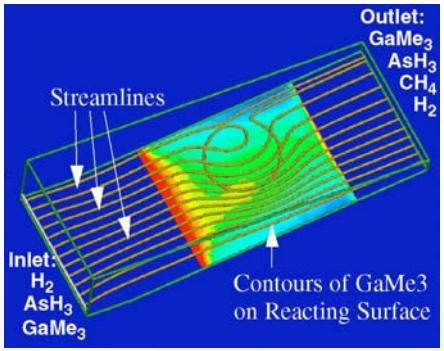


Reliability: b>2
10x robustness

Fully integrated, adaptive approach is more accurate, less expensive, more reliable, and more convenient

DAKOTA Optimization Toolkit includes diverse capabilities.





Goal: **answer fundamental engineering questions**

- **What is the best design?**
- **How safe is it?**
- **How much confidence in my answer?**

Challenges

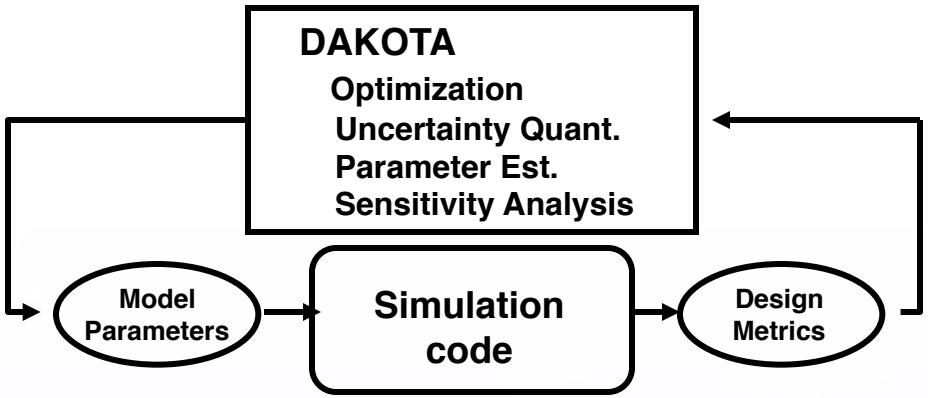
- **Reuse tools and interfaces, leverage commonalities ⇒ software**
- **Nonsmooth/discontinuous/multimodal, expensive, mixed variables, unreliable gradients, simulation failures ⇒ algorithm R&D**
- **ASCI-scale applications & architectures ⇒ scalable parallelism**

Impact

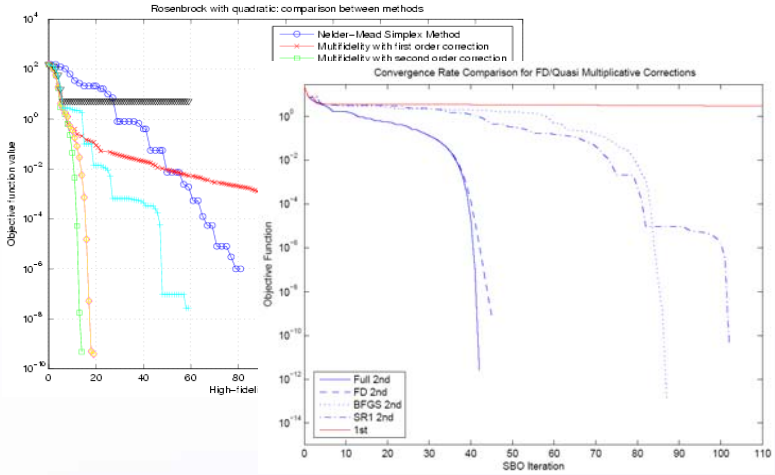
- **DOE: Tri-lab tool, broad application deployment**
- **External: WFO partners, GNU GPL (>3000 download registrations)**

Sandia National Laboratories

Optimization and Uncertainty Quantification with Dakota has enabled many advances.

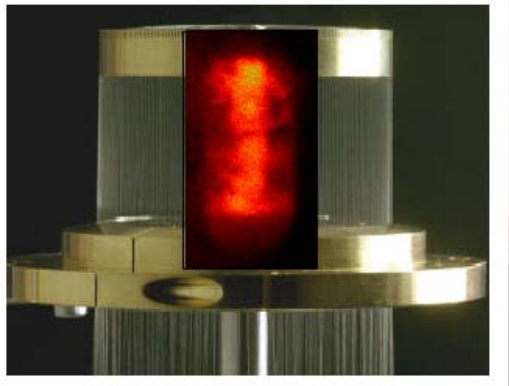


Surrogate-based



Z pinch flyer plate

(useful EOS, world record velocities)



JSF shape optimization

(max range & min yaw @ separation, validated)

DAKOTA

Sandia National Laboratories

A variety of software tools are deployed in computational UQ efforts.

- **DAKOTA is a major tool developed by Sandia and is available through open release.**
- **Many other tools are commercial and/or commonly available:**
  - Minitab statistics package
  - JMP statistics package
  - Mathematica
  - Matlab with Statistics Toolbox
  - R or S+ language
  - Simlab
  - Excel add-ins (@Risk, Crystal Ball)
  - Others…
- **For example, when used in sensitivity analysis …**

What is Sensitivity Analysis (SA)?

SA seeks to quantify at the influence of the **uncertainty** in the input on the **uncertainty** of the output

SA strives to help answer the question:

> *"How important are the individual elements of input* x *with respect to the uncertainty in output* $y(x)$*?"*

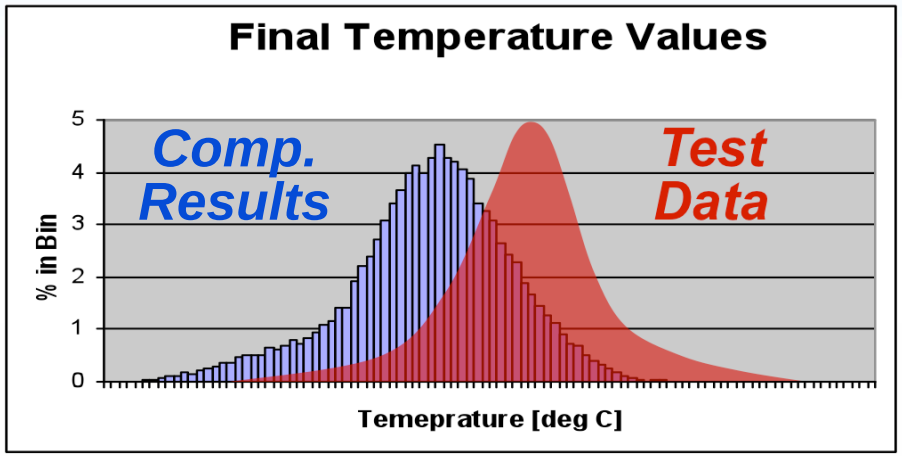
SA can be used to:

- Rank input parameters in term of their importance relative to the uncertainty in the output;
- Support verification and validation activities;
- Drive, as part of an iterative process, uncertainty quantification (UQ) analyses towards the input parameters that really matter.

SA is typically the starting point for a more complete UQ.

SA is not a replacement for full UQ or V&V.

In some sense, much of our work seeks to create and understand diagrams like this:



- **This diagram is a simple illustration of:**
  - A large number (“ensemble”) of complex calculations representing computational UQ, given as an estimated probability density;
  - A representation of observational data with characterized uncertainty, given as an estimated probability density;
  - A rigorous means of comparing the two types of information;
  - A rigorous means of inferring model “quality” (“accuracy,” “credibility”) from the comparison;
  - Typically undertaken for purposes of understanding the limits of applicability of the model.

What are Verification and Validation?

Verification seeks to answer the question:

"Is my computational model, as instantiated in software, solving the governing equations correctly?"

- Verification is primarily about math and computer science.
- Verification comes in different flavors:
  - Software verification
  - Code verification
  - Calculation verification

Validation seeks to answer the question:

"Is my computational model, as instantiated in software, solving the proper governing equations?"

- Validation is primarily about modeling and "physics."
- Validation *necessarily* involves data.
- Validation intersects with other difficult problems:
  - Calibration
  - Uncertainty Quantification
  - Sensitivity Analysis

② Sandia has made key contributions to the development of HPC technology.


CM-2


nCUBE-2


iPSC-860


Paragon


ASCI Red


Cplant


Red Storm

1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009

Gordon Bell Prize

R&D 100
Parallel Software

Patent
Meshing

R&D 100
Dense Solvers

Gordon Bell Prize

R&D 100
Storage

World Record
Teraflops

Gordon Bell Prize

R&D 100
Xyce

R&D 100
Allocator

Patent
Data Mining

R&D 100
Signal Processing

SC96 Gold Medal
Networking

Mannheim
SuParCup

R&D 100
Catamount

Karp Challenge

World Record
281 GFlops

R&D 100
Salvo

R&D 100
Meshing

Patent
Parallel Software

World Record
143 GFlops

Patent
Partitioning

R&D 100
Trilinos

Patent
Paving

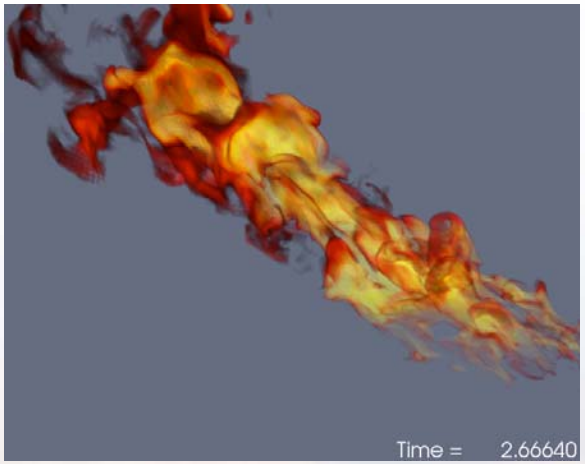
R&D 100
Aztec

Fernbach
Award

Sandia National Laboratories

Red Storm is Sandia's flagship HPC platform.

- **True MPP, designed to be a single system.**
  - Full 3-D mesh interconnect
  - 12,960 compute nodes (Dual core AMD Opterons @ 2.4 GHz)
  - 35 Terabytes of memory
  - 400 Terabytes of disk storage
  - Tenths-of-Pflop peak performance
- **Sandia contributions included:**
  - Helped design interconnect
  - Developed compute node operating system based on Sandia's light weight kernel technology
- **Diverse set of NNSA problems have been solved on Red Storm.**



Fire simulations are being used to certify new test facility and will be used to analyze weapons safety issues.

“cielo” is recent Sandia collaborative HPC effort.

- **cielo represents a major collaboration between Los Alamos and Sandia to jointly design, build, and site a Petaflop computer at LANL for NNSA.**

cielo

NNSA New Mexico Alliance for Computing at Extreme Scale

1.03 PFlops peak entering service

cielo by the numbers…

Design Points

	Phase 1	Phase 2	Cielito
# of Cabinets	72	96	1
# of Service Nodes	208	272	14
# of Compute Nodes	6,704*	8,944*	68
# of Visualization Nodes	(376)	(376)	(4)
# of Compute Cores	107,264	143,104	1,088
Peak Memory BW	572 TB/s	763 TB/s	5.8 TB/s
Memory Capacity per Core	2 GB (4 GB)	2 GB (4 GB)	2 GB (4 GB)
Compute Memory Capacity	226.6 TB	298.2 TB	2.3 TB
Peak Compute FLOPS	1.03 PF	1.37 PF	10.4 TF
Sustained PFS BW	> 160 GB/s		TBD
System Power	< 3.9 MW	< 4.4 MW	
Full System Job MTBI	> 25 hours		
System MTBI	> 200 hours		
* Total compute nodes including Viz nodes and nodes allocated for other services			

Current Status

- **Target was 6x performance improvement over prior platform (ASC Purple)**
- **Achieved 9.6x improvement over a range of target applications!**
- **Job MTBI ~ 15 hours and improving**
- **System MTBI > 200 hours (achieved)**

Sandia National Laboratories

What's next in HPC: the DOE Exascale Initiative.

The DOE Exascale Initiative is a national-scale computing strategy that teams:

- DOE National Nuclear Security Administration (NNSA) and DOE Advanced Scientific Computing (ASC) program, together with
- DOE Office of Science (SC) Advanced Scientific Computing Research (ASCR) program.

Intent: to develop the national vision for next generation high performance computing applications and architectures.

Goals:

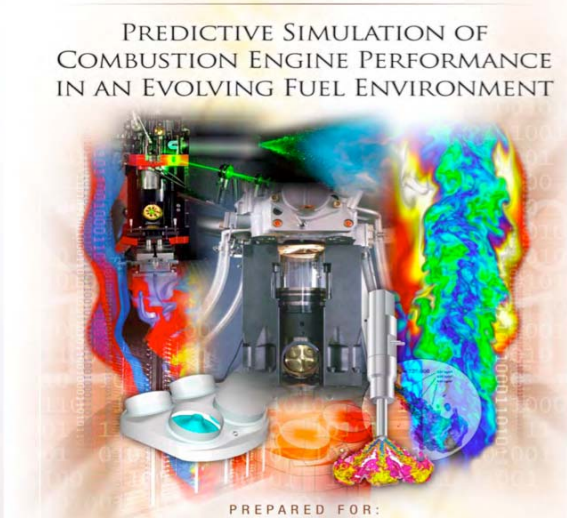
- **2015**: **200 PF/s peak, 0.5 – 7 TF/node, $5\text{´}10^3 – 5\text{´}10^4$ nodes, 15 MW**
- **2018**: **1 EF/s peak, 1 – 10 TF/node, $1\text{´}10^5 – 1\text{´}10^6$ nodes, 20 MW**

Sandia National Laboratories

Sandia is influencing Exascale discussions at the national level.

- **Helped lead development of DOE's Exascale technology roadmap**
  - Basis for presentations to highest levels of DOE management (Chu, D'Agostino, Koonin)
- **Influenced key elements of strategy, especially co-design**
  - 2 plenary presentations on co-design
  - Widely acknowledged by community as critical to success
- **Helped establish combustion as one of 4 key DOE Exascale applications**
- **Sandia lead the Structural Simulation Toolkit**
- **Sandia helps leads in architecture-aware algorithms**

Sandia was heavily involved in the Exascale workshops



Potential Exascale system architectures have been scoped.

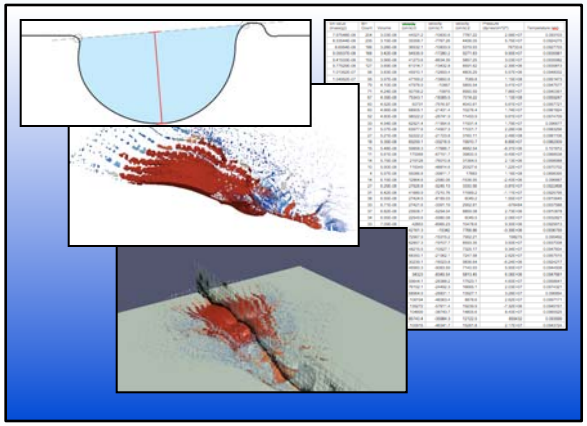
System attributes	2010	"2015"		"2018"	
System peak	2 Peta	200 Petaflop/sec		1 Exaflop/sec	
Power	6 MW	15 MW		20 MW	
System memory	0.3 PB	5 PB		32–64 PB	
Node performance	125 GF	0.5 TF	7 TF	1 TF	10 TF
Node memory BW	25 GB/s	0.1 TB/sec	1 TB/sec	0.4 TB/sec	4 TB/sec
Node concurrency	12	*O*(100)	*O*(1,000)	*O*(1,000)	*O*(10,000)
System size (nodes)	18,700	50,000	5,000	1,000,000	100,000
Total Node Interconnect BW	1.5 GB/s	20 GB/sec		200 GB/sec	
MTTI	~ days	*O*(1 day)		*O*(1 day)	

Sandia National Laboratories

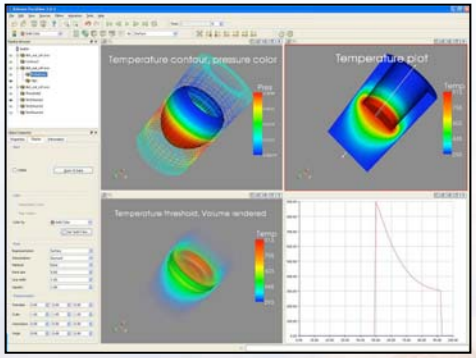
③ Sandia's scalable analysis tools promote insight into complex data.

Insight requires powerful tools to help customers analyze and understand their data. By providing a continuum of technologies — applications, scripting capabilities, and coupled analysis libraries — we provide the means to connect simulations, test results, and collaborative knowledge.

- **Scalable Analysis Applications**
  - **ParaView, Prism, ImageDrill, Threatview, VxInsight, Titan, OverView**
- **Sciviz and Infoviz tools for analysis of complex simulation data**
  - Families of runs, complex feature recognition and analysis
  - Coupled SciViz and InfoViz
- **Coupled Analysis Tools**
  - In concert with a running simulation, produce a rich set of output data, including images, statistics, features (fragments), and meshes
- **Flexible scripting tools for large data**
  - Promotes rapid development of project-driven solutions for large data

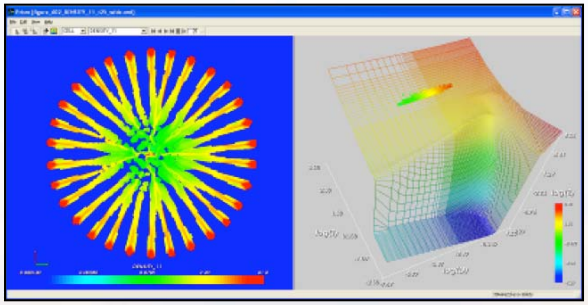


Examples of rich data output – feature detection, computed values, images and meshes – computed during a running simulation with coupled analysis capabilities.

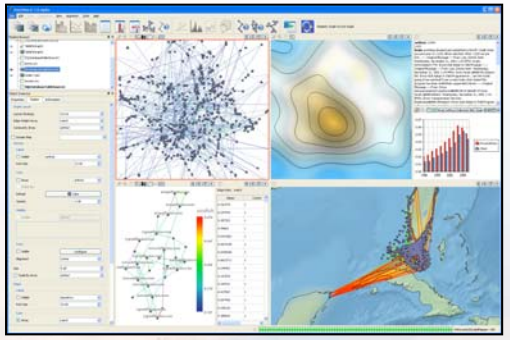


ParaView: an open source, scalable scientific analysis application, provides a flexible platform for scalable tool development.

http://www.paraview.org/



Prism: a ParaView-based application that couples scientific visualization with abstract information visualization in a phase state surface. Material can be explored in both views at the same time.

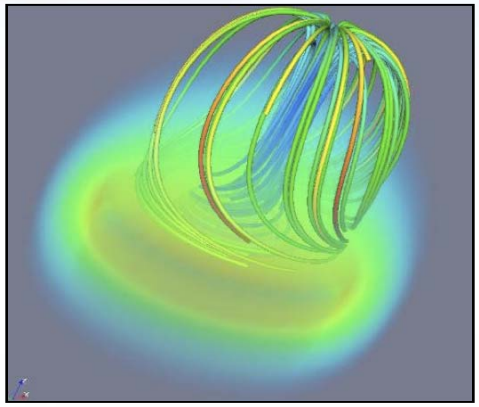


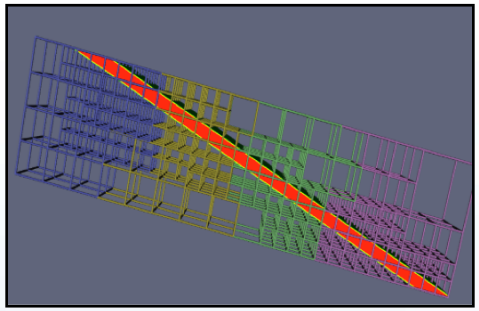
The same technology is now being expanded to information visualization techniques, promoting understanding of many types of complex data.

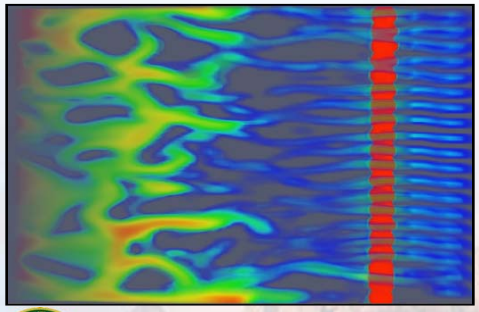
http://titan.sandia.gov/

Sandia National Laboratories

Visualization and analysis of M&S is essential to our toolbox.







Scientific Visualization

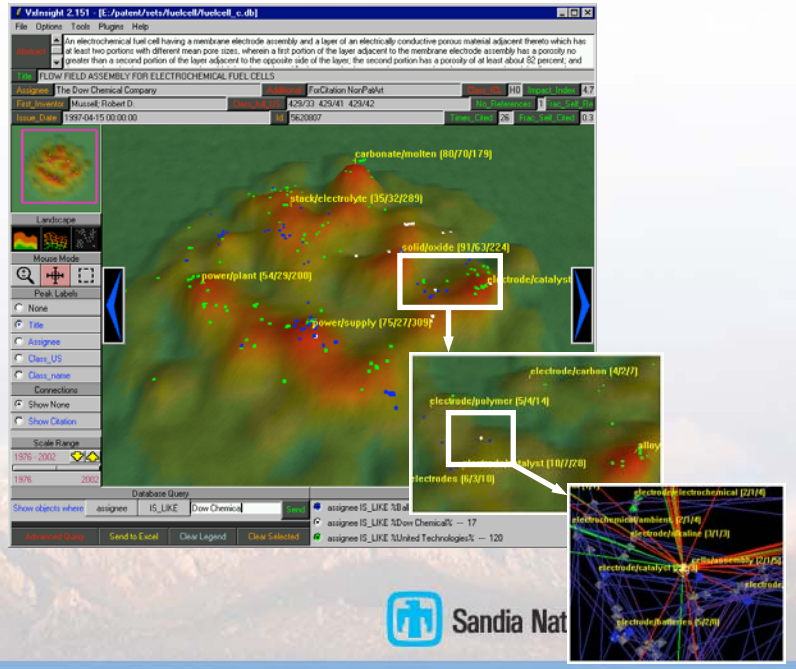
ParaView

- **Parallel visualization tool**
- **Used to visualizes Sandia's largest FEA simulations.**
- **Software base is very agile:**
  - Open-source visualization tool
  - Allows rapid reconfiguration and application to new areas

Informatics Visualization

Threatview, VxInsight, Titan, OverView

- **Displays the structure of network information.**
- **Allows the analyst to explore information at many levels.**



Sandia Nat

④ Optimization of Large-Scale Heterogeneous System-of-Systems Models

Bill Hart, Scott Collis in 1400

Heterogeneous system of systems (HSoS) integrate subsystem models to describe large, complex systems.

Goal: Efficiently analyze uncertainties in large-scale models of complex man-made systems

Analysis of Policy Questions:

- What resource and infrastructure bottlenecks can critically limit growth of U.S. energy?
- How might shifts in world climate impact the availability and reliability of energy?
- What mix of energy resources meets environmental goals and minimizes costs?
- Where would technological improvements have the most impact?

Technical Focus in Four Research Areas:

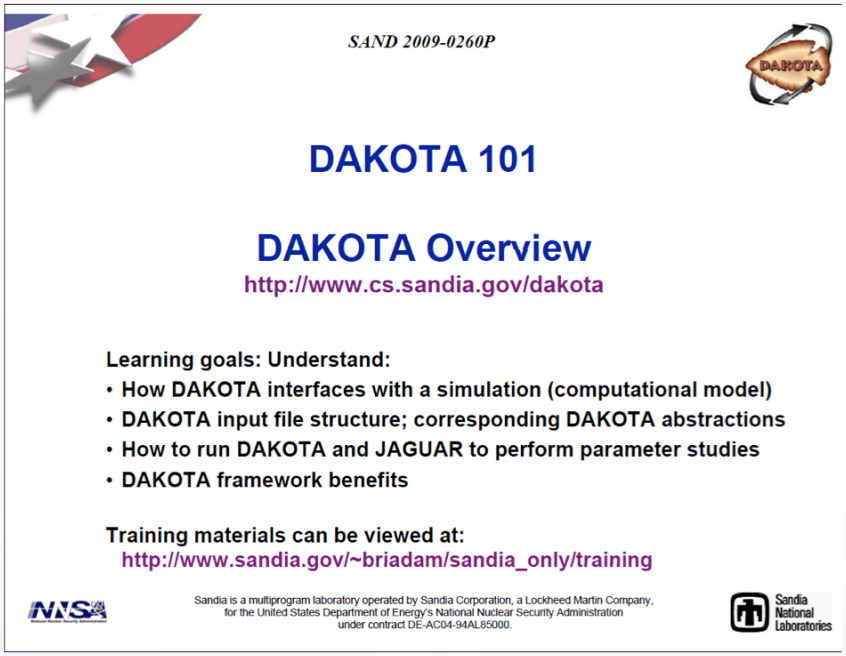
- ***Multi-stage stochastic optimization*** *with recourse* to plan and manage system evolution through time, given uncertain information about the future.
- ***Risk management*** to identify system parameters that are insensitive to data uncertainties.
- ***Multi-objective optimization*** to assess tradeoffs between HSoS performance criteria, including risk, cost, and system performance.
- ***Model consistency methods*** to find critical modeling uncertainties and quantify model confidence.

Major Goals/Milestones:

- Develop parallel multi-stage optimizers with asynchronous communication
- Implement scalable risk-management optimization strategies
- Develop model consistency metrics to guide HSoS model refinement
- Develop multi-objective optimization strategy with uncertain objectives
- ***Analysis of large-scale HSoS applications****:*
  - Future energy infrastructure
  - Multi-platform intelligence collection systems

Sandia National Laboratories

Sandia supports its tools with training.



SAND 2009-0260P

DAKOTA 101

DAKOTA Overview

http://www.cs.sandia.gov/dakota

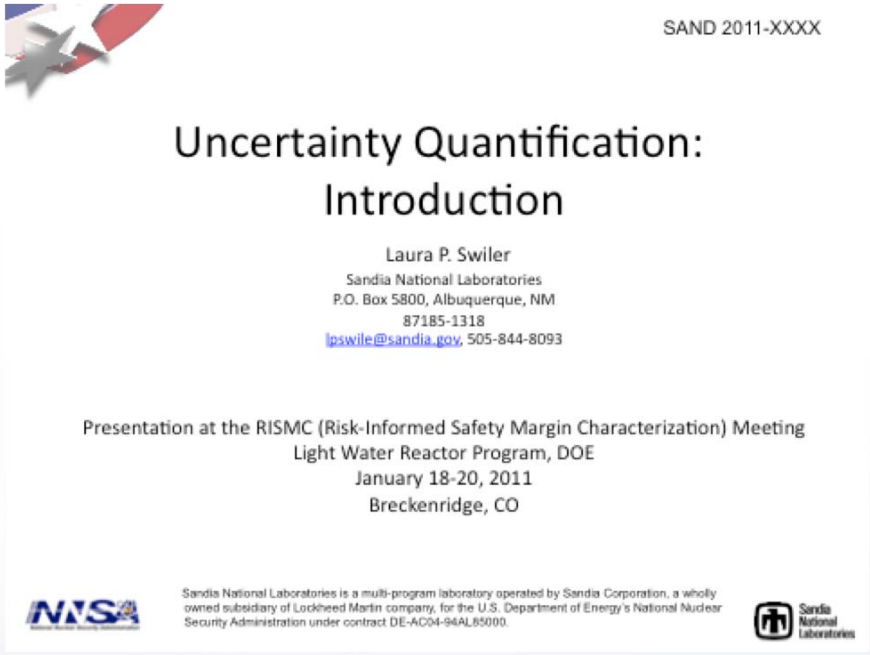
Learning goals: Understand:

- **How DAKOTA interfaces with a simulation (computational model)**
- **DAKOTA input file structure; corresponding DAKOTA abstractions**
- **How to run DAKOTA and JAGUAR to perform parameter studies**
- **DAKOTA framework benefits**

Training materials can be viewed at:
http://www.sandia.gov/~briadam/sandia_only/training

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories



SAND 2011-XXXX

Uncertainty Quantification: Introduction

Laura P. Swiler
Sandia National Laboratories
P.O. Box 5800, Albuquerque, NM
87185-1318
lpswile@sandia.gov, 505-844-8093

Presentation at the RISMC (Risk-Informed Safety Margin Characterization) Meeting
Light Water Reactor Program, DOE
January 18-20, 2011
Breckenridge, CO

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

- **Education about domain knowledge and tool use is an important goal of 1400.**
- **Training for viz tools is also part of our portfolio.**

Sandia National Laboratories

There are specific areas for Sandia and UVM to develop interactions.

- There appear to be clear areas for interaction between Sandia Computing Research Center—Org. 1400—and UVM in areas of mutual interest that support the goals of this activity.

	SNL Training	*SNL® UVM*	*UVM®SNL*
① Uncertainty Quantification	✓	✓	✓
② High Performance Computing	?	?	✓
③ Visualization: SciViz & InfoViz	✓	✓	✓
④ Optimization of HSoS	?	✓	✓

SNL algorithms capabilities enable much continued M&S transformation.

- **Trilinos:** HPC solution components
  - Compatible space-time discretizations
  - Linear & nonlinear solvers
  - Partitioning and dynamic load balancing
  - Automatic differentiation
  - Optimization (fully coupled)

Full integration

- **Dakota:** Risk informed decision making
  - Optimization (multiparallel, surrogate,..)
  - UQ (aleatory & epistemic)
  - New sparse-collocation methods…
  - Optimization under uncertainty
  - Rapid deployment of new algorithms
- Extensive algorithms R&D foundation…

Ready to use optimization and UQ methods

Trilinos is a principal software platform for scalable algorithms.

- Cutting edge algorithmic research on:
  - Sparse, distributed, parallel linear algebra,
  - Iterative and direct parallel linear solvers,
  - Iterative, parallel eigensolvers,
  - Multilevel parallel preconditioners,
  - Load-balancing, AD, and more…
- State-of-the-art solver framework:
  - Generic, object-oriented programming
  - Extensible, scalable design
  - Common core for linear algebra with abstract solver API
  - SQE infrastructure with requirements and best practices for SQA
- Impact
  - Available in most ASC applications codes
  - Impacts several CRADA projects
  - Used by all major DOE labs
  - *>2300 registered users*
- 2004 R&D 100 award
- IEEE 2004 HPC Software Challenge Award